

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 20:46:52 ; Search time 8673 Seconds
(without alignments)
17519.399 Million cell updates/sec

Title: US-09-918-187-3
Perfect score: 5221
Sequence: 1 ataaaggggctgaggaaa.....aatctaaaaaaaaaaaaa 5221

Scoring table: IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 993650

Minimum DB seq length: 8
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 39 | 0.7 | 39 | 6 AR003340 | AR003340 Sequence |
| 2 | 33.6 | 0.6 | 86 | 3 AY009729 | AY009729 Boophilus |
| 3 | 32.4 | 0.6 | 90 | 4 OAAJ5003 | AJ005003 Ovis ari |
| 4 | 31.6 | 0.6 | 96 | 9 HUMDINUJ | L21846 Human dinuc |
| 5 | 31.2 | 0.6 | 90 | 10 RN012243 | U12243 Rattus norv |
| 6 | 31 | 0.6 | 69 | 4 AF339940 | AF339940 Puma conc |
| 7 | 31 | 0.6 | 69 | 4 AF339942 | AF339942 Puma conc |
| 8 | 30.6 | 0.6 | 83 | 5 AF420492 | AF420492 Salmo sal |
| 9 | 30.4 | 0.6 | 71 | 4 AF055530 | AF055530 Didelphis |
| 10 | 30.4 | 0.6 | 73 | 4 AF339937 | AF339937 Puma conc |
| 11 | 30.2 | 0.6 | 88 | 3 AY009730 | AY009730 Boophilus |
| 12 | 30 | 0.6 | 73 | 6 AX340088 | AX340088 Sequence |
| 13 | 30 | 0.6 | 73 | 11 AF042300 | AF042300 Homo sapi |
| 14 | 29.6 | 0.6 | 92 | 11 G68120 | G68120 D7S3168 (CR |
| 15 | 29.6 | 0.6 | 93 | 6 I05524 | I05524 Sequence 11 |
| 16 | 29.6 | 0.6 | 95 | 4 AF126190 | AF126190 Cervus da |
| 17 | 29.4 | 0.6 | 73 | 6 AX397053 | AX397053 Sequence |
| 18 | 29.4 | 0.6 | 95 | 11 AU026788 | AU026788 Rattus no |
| 19 | 29.2 | 0.6 | 97 | 8 AB055133 | AB055133 Polygonum |
| 20 | 29 | 0.6 | 97 | 4 AF213233 | AF213233 Bison bis |
| 21 | 29 | 0.6 | 99 | 11 AU049997 | AU049997 Rattus no |
| 22 | 28.8 | 0.6 | 50 | 6 I31455 | I31455 Sequence 36 |
| 23 | 28.8 | 0.6 | 65 | 6 I31204 | I31204 Sequence 11 |
| 24 | 28.8 | 0.6 | 65 | 6 I31509 | I31509 Sequence 42 |
| 25 | 28.8 | 0.6 | 75 | 11 AF042301 | AF042301 Homo sapi |
| 26 | 28.8 | 0.6 | 77 | 11 AF042302 | AF042302 Homo sapi |
| 27 | 28.8 | 0.6 | 91 | 3 AY009760 | AY009760 Boophilus |
| 28 | 28.6 | 0.5 | 87 | 9 HUML | L23536 Human chrom |
| 29 | 28.6 | 0.5 | 87 | 11 G31698 | G31698 swSS1885 Er |
| 30 | 28.6 | 0.5 | 89 | 11 AU027741 | AU027741 Rattus no |
| 31 | 28.6 | 0.5 | 97 | 4 MMA309850 | AJ309850 Martes ma |
| 32 | 28.4 | 0.5 | 73 | 11 AF042303 | AF042303 Homo sapi |
| 33 | 28.4 | 0.5 | 91 | 10 MMDNDS20 | X55238 M.musculus |
| 34 | 28.4 | 0.5 | 93 | 4 AF213232 | AF213232 Bison bis |
| 35 | 28.2 | 0.5 | 51 | 6 I31425 | I31425 Sequence 33 |
| 36 | 28.2 | 0.5 | 60 | 3 DRORS2 | M35655 D.melanogas |
| 37 | 28.2 | 0.5 | 62 | 4 AF185112 | AF185112 Lasiorhin |
| 38 | 28.2 | 0.5 | 69 | 11 AF042298 | AF042298 Homo sapi |
| 39 | 28.2 | 0.5 | 85 | 10 MMDNDS16 | X55234 M.musculus |
| 40 | 28.2 | 0.5 | 86 | 6 AX203294 | AX203294 Sequence |
| 41 | 28.2 | 0.5 | 92 | 5 AF019993 | AF019993 Clupea pa |
| 42 | 28.2 | 0.5 | 95 | 4 AF213195 | AF213195 Bison bis |
| 43 | 28 | 0.5 | 79 | 3 AY009762 | AY009762 Boophilus |
| 44 | 28 | 0.5 | 83 | 11 AF042305 | AF042305 Homo sapi |
| 45 | 28 | 0.5 | 87 | 11 HUMUT792A | L31205 Human STS u |

ALIGNMENTS

RESULT 1
AR003340
LOCUS AR003340 39 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 62 from patent US 5744300.
ACCESSION AR003340
VERSION AR003340.1 GI:3964599
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 39)
AUTHORS Linskens,M.H.K., Hirsch,K.S., Villeponteau,B., Feng,J., Funk,W. and West,M.David.
TITLE Methods and reagents for the identification and regulation of senescence-related genes

JOURNAL Patent: US 5744300-A 62 28-APR-1998;
FEATURES location/Qualifiers
Source 1..39 /organism="unknown"
BASE COUNT 11 a 7 c 6 g 15 t
ORIGIN
Query Match 0.7%; Score 39; DB 6; Length 39;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1412 AGATGATGATGTTAACCATTCCAGTACAGTATCTTTT 1450
DB 1 AGATGATGATGTTAACCATTCCAGTACAGTATCTTTT 39
RESULT 2
AY009729 86 bp DNA linear INV 11-APR-2001
LOCUS Boophilus microplus clone PNC 75 isolate Eacham2R microsatellite
DEFINITION sequence.
ACCESSION AY009729
VERSION AY009729.1 GI:13603305
KEYWORDS
SOURCE Boophilus microplus.
ORGANISM Boophilus microplus.
REFERENCE 1 (bases 1 to 86)
AUTHORS Chigagure,N.N. and Barker,S.C.
TITLE Microsatellite DNA population genetic analysis of Boophilus
microplus (Acari:Ixodidae)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 86)
AUTHORS Chigagure,N.N. and Barker,S.C.
TITLE Direct Submission
JOURNAL Submitted (11-OCT-2000) Parasitology and Microbiology, The
University of Queensland, St. Lucia Campus, Brisbane, Queensland
QLD 4072, Australia
FEATURES location/Qualifiers
Source 1..86
/organism="Boophilus microplus"
/isolate="Eacham2R"
/db_xref="taxon:6941"
/clone="PNC 75"
/country="Australia"
repeat_region
/note="microsatellite"
/rpt_type=tandem
/rpt_unit=ca
BASE COUNT 7 a 11 c 40 g 28 t
ORIGIN
Query Match 0.6%; Score 33.6; DB 3; Length 86;
Best Local Similarity 70.3%; Pred. No. 2.6e+03;
Matches 45; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
OY 4052 ATGGCACCCTCAGGCTGAGGGCCCCAATGATGTGTGGCTGTGGGTGGGTGGAGTGTG 4111
DB 12 ATGACACTGTAGTGGGGCGGCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 71
OY 4112 TCTG 4115
DB 72 TGTG 75
RESULT 3
OAAJ5003/c 90 bp DNA linear MAM 01-OCT-1998
LOCUS Ovis arles gene encoding brain-type ribonuclease, 3'UTR.
DEFINITION
ACCESSION AJ005003
VERSION AJ005003.1 GI:3005979
KEYWORDS ribonuclease.

SOURCE Ovis arles.
ORGANISM Ovis arles
REFERENCE 1 (bases 1 to 90)
AUTHORS Furia,A.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-1998) Furia A., Department of Organic and
Biological Chemistry, University of Naples Federico II, Via
Mazzocanone 16, 80134, ITALY
2 (bases 1 to 90)
REFERENCE 2 (bases 1 to 90)
AUTHORS Breukelman,H.J., van der Munnik,N., Kleineldam,R.G., Furia,A. and
Beintema,J.J.
TITLE Secretory ribonuclease genes and pseudogenes in true ruminants
JOURNAL Gene 212 (2), 259-268 (1998)
MEDLINE 98278842
PUBMED 9611269
FEATURES location/Qualifiers
Source 1..90
/organism="Ovis arles"
/db_xref="taxon:9940"
BASE COUNT 3'UTR 32 a 29 c 10 g 19 t
ORIGIN
Query Match 0.6%; Score 32.4; DB 4; Length 90;
Best Local Similarity 72.4%; Pred. No. 5.4e+03;
Matches 42; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
OY 4078 TGTATGTGTGGCTGTGGGTGGGTGGAGTGTGTCTGTGAGTAAAGAACAGATT 4135
DB 75 TATACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAATAAGAAACCTACTT 18
RESULT 4
HUMDINUJ 96 bp DNA linear PRI 27-JUL-1994
LOCUS Human dinucleotide repeat polymorphism from 16q24.2-q24.3 DNA
DEFINITION sequence.
ACCESSION L21846
VERSION L21846.1 GI:516079
KEYWORDS dinucleotide repeat; dinucleotide repeat polymorphism; repeat
polymorphism.
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 96)
AUTHORS Shen,Y.
TITLE Five dinucleotide repeat polymorphisms on human chromosome
16q24.2-q24.3
JOURNAL Hum. Mol. Genet. (1993) In press
FEATURES location/Qualifiers
Source 1..96
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 15 a 3 c 42 g 36 t
ORIGIN
Query Match 0.6%; Score 31.6; DB 9; Length 96;
Best Local Similarity 62.8%; Pred. No. 8.8e+03;
Matches 49; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
OY 4045 AGCTGATGACACCTCAGCTGAGGGCCCCAATGATGTGTGGCTGTGGGTGGGTGG 4104
DB 17 AGGAGCAGCGAATGAGAGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 76
OY 4105 GAGTGTCTGTGAGTA 4122
DB 77 GTGTGTGTGTGTGTGA 94

[illegible]

```

FEATURES
134 Cheatham, Blacksburg, VA 24061-0321, USA
Location/Qualifiers
1..69
/organism="Puma concolor"
/db_xref="taxon:9696"
/chromosome="A1"
/map="locus FCA090"
/clone="pc0553"
1..69
repeat_region
/note="microsatellite"
/rpt_type=tandem
BASE COUNT 33 a 24 c 6 g 6 t
ORIGIN
Query Match 0.6%; Score 31; DB 4; Length 69;
Best Local Similarity 72.7%; Pred. No. 1.2e+04;
Matches 40; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Db 57 TCATACCTTAGAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCT 3

RESULT 7
AF339942/c 69 bp DNA linear MAM 12-JUN-2001
DEFINITION Puma concolor clone Pco564 locus FCA090 microsatellite sequence.
ACCESSION AF339942
VERSION AF339942.1 GI:14348697
KEYWORDS
SOURCE
ORGANISM
Puma concolor.
Puma concolor
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma.
REFERENCE
1 (bases 1 to 69)
Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J.
Patterns of size homoplasy at 10 microsatellite loci in pumas (Puma
concolor)
Mol. Biol. Evol. 18 (6), 1151-1156 (2001)
JOURNAL
MEDLINE 21265424
PUBMED 11371606
2 (bases 1 to 69)
Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J.
Direct Submission
Submitted (24-JAN-2001) Fisheries and Wildlife Sciences, VPI&SU,
134 Cheatham, Blacksburg, VA 24061-0321, USA
JOURNAL
FEATURES
source
location/Qualifiers
1..69
/organism="Puma concolor"
/db_xref="taxon:9696"
/chromosome="A1"
/map="locus FCA090"
/clone="pc0564"
1..69
repeat_region
/note="microsatellite"
/rpt_type=tandem
BASE COUNT 33 a 24 c 6 g 6 t
ORIGIN
Query Match 0.6%; Score 31; DB 4; Length 69;
Best Local Similarity 72.7%; Pred. No. 1.2e+04;
Matches 40; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Db 57 TCATACCTTAGAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCT 3

RESULT 8
AF420492 83 bp DNA linear VRT 03-OCT-2001
DEFINITION Salmo salar clone SsalNWIG microsatellite sequence.
ACCESSION AF420492

```

```

VERSION      AF420492.1  GI:15866784
KEYWORDS
SOURCE       Salmo salar.
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei;
              Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE    1 (bases 1 to 83)
              Powell,R.
              Microsatellite DNA sequences from the Atlantic salmon
TITLE        Unpublished
JOURNAL      2 (bases 1 to 83)
AUTHORS      Powell,R.
REFERENCE    Direct Submission
              Submitted (18-SEP-2001) Department of Microbiology, National
              University of Ireland, Galway (NUIG), Galway, Ireland
FEATURES
  source     1..83
              /organism="Salmo salar"
              /db_xref="taxon:8030"
              /clone="SaalNUIG"
              19..64
              /note="microsatellite"
              /rpt_type=tandem
              /rpt_unit=ca
BASE COUNT   33 a 35 c 11 g 4 t
ORIGIN
Query Match          0.6%; Score 30.4; DB 5; Length 83;
Best Local Similarity 73.6%; Pred. No. 1.6e+04;
Matches 39; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

      .
      repeat_region
      /note="microsatellite"
      /rpt_type=tandem
      /rpt_unit=ca
BASE COUNT   33 a 35 c 11 g 4 t
ORIGIN
Query Match          0.6%; Score 30.4; DB 4; Length 71;
Best Local Similarity 67.2%; Pred. No. 1.8e+04;
Matches 43; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY  4693 ACTTCGTGTGACCCCTGGGCAAGTCACTTAAGTAAAGTGCCTCAGTTTCCTCTCTGT 4752
      ||| ||||| ||||| ||||| ||||| ||| | | | | | ||| ||| ||| ||| |||
      5 ACTACCTGTGTGACCTTGGGCAAGTCACTTACCCCTGTGGCCTCAGTTTCTCATCTGT 64

RESULT 9
AF055530      71 bp      DNA      linear      MAM 16-OCT-2001
LOCUS         AF055530
DEFINITION    Didelphis virginiana isolate 051 retroposon CORE-SINE Mar-1
sequence.
ACCESSION     AF055530
VERSION       AF055530.1  GI:4511910
KEYWORDS
SOURCE        Didelphis virginiana.
              Didelphis virginiana
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
REFERENCE    1 (bases 1 to 71)
              Gilbert,N. and Labuda,D.
              Evolutionary inventions and continuity of CORE-SINES in mammals
              J. Mol. Biol. 298 (3), 365-377 (2000)
MEDLINE       20237720
PUBMED        10772856
FEATURES
  source     1..71
              Location/Qualifiers
              /organism="Didelphis virginiana"
              /isolate="051"
              /db_xref="taxon:9267"
              1..71
              /rpt_family="retroposon CORE-SINE Mar-1"
              /rpt_type=dispersed
BASE COUNT   14 a 19 c 15 g 23 t
ORIGIN
Query Match          0.6%; Score 30.4; DB 4; Length 71;
Best Local Similarity 67.2%; Pred. No. 1.8e+04;
Matches 43; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```

| | | | |
|----|------|-----|------|
| OY | 4753 | TAA | 4756 |
| | | | |
| Db | 65 | CAG | 68 |

```
RESULT 10  
LOCUS AF339937/c 73 bp DNA linear MAM 12-JUN-2001  
DEFINITION Puma concolor clone Fco036 locus FCA090 microsatellite sequence.  
ACCESSION AF339937  
VERSION AF339937.1 GI:14348692  
KEYWORDS  
SOURCE Puma concolor.  
ORGANISM Puma concolor  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma.  
REFERENCE  
AUTHORS 1 (bases 1 to 73)  
TITLE Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J.  
Patterns of size homoplasmy at 10 microsatellite loci in pumas (Puma concolor)  
JOURNAL Mol. Biol. Evol. 18 (6), 1151-1156 (2001)  
MEDLINE 21265424  
PubMed 11371606  
REFERENCE 2 (bases 1 to 73)  
AUTHORS Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J.  
TITLE Direct Submission  
JOURNAL Submitted (24-JAN-2001) Fisheries and Wildlife sciences, VPI&SU,  
134 Cheatham, Blacksburg, VA 24061-0321, USA  
  
FEATURES  
source  
    1..73  
        /organism="Puma concolor"  
        /db_xref="taxon:9696"  
        /chromosome="A1"  
        /map="locus FCA090"  
        /clone="Fco036"  
    1..73  
        /note="microsatellite"  
repeat_region  
    /rpt_type=tandem  
BASE COUNT 35 a 26 c 6 g 6 t  
ORIGIN  
  
Query Match 0.6%; Score 30.4; DB 4; length 73;  
Best local Similarity 71.4%; Pred. NO. 1.8e+04;  
Matches 40; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

| | | | |
|----|------|---|------|
| OY | 4060 | TCAGCGTGAGGGCCCCCAATGATGTGTGGCTGTGGCTGTGGTGGAGTGTCTGC | 4115 |
| | | | |
| Db | 61 | TCATACTTAGAGCTGCTGC | 6 |

```
RESULT 11  
LOCUS AY009730 88 bp DNA linear INV 11-APR-2001  
DEFINITION Boophilus microplus clone PNC 75 isolate Eacham3R microsatellite  
sequence.  
ACCESSION AY009730  
VERSION AY009730.1 GI:13603306  
KEYWORDS  
SOURCE Boophilus microplus.  
ORGANISM Boophilus microplus  
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
Parasitiformes; Ixodida; Ixodidae; Boophilus.  
REFERENCE 1 (bases 1 to 88)  
AUTHORS Chigagure,N.N. and Barker,S.C.  
TITLE Microsatellite DNA population genetic analysis of Boophilus  
microplus (Acari:Ixodidae)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 88)  
AUTHORS Chigagure,N.N. and Barker,S.C.  
TITLE Direct Submission  
JOURNAL Submitted (11-OCT-2000) Parasitology and Microbiology, The  
University of Queensland, St. Lucia Campus, Brisbane, Queensland
```

| | |
|-----------------------|--|
| FEATURES | OLD 4072, Australia |
| source | Location/Qualifiers |
| | 1. .88 |
| | /organism="Boophilus microplus" |
| | /isolate="Eacham3R" |
| | /db_xref="taxon:6941" |
| | /clone="PNC 75" |
| | /country="Australia" |
| | 1. .88 |
| repeat_region | /note="microsatellite" |
| | /rpt_type=tandem |
| | /rpt_unit=ca |
| BASE COUNT | 9 a 8 c 35 g 36 t |
| ORIGIN | |
| Query Match | 0.6%; Score 30.2; DB 3; Length 88; |
| Best Local Similarity | 62.7%; Pred. No. 2e+04; |
| Matches | 47; Conservative 0; Mismatches 28; Indels 0; Gaps 0; |
| QY | 4038 CTACATAAGGCTGCATGGCACTCAGGCTGAGGGCCCAATGATGTGTGGCTGTGGCTG 4097 |
| Db | 5 CGACATGAAGTGGCACTGCTACTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 64 |
| QY | 4098 TGGGTGGGAGTGTGT 4112 |
| Db | 65 TGTGTGTGTGTGTGT 79 |
| RESULT 12 | |
| LOCUS | AX340088 63 bp DNA linear PAT 10-JAN-2002 |
| DEFINITION | Sequence 335 from Patent WO0196388. |
| ACCESSION | AX340088 |
| VERSION | AX340088.1 GI:18136069 |
| KEYWORDS | human. |
| SOURCE | |
| ORGANISM | Homo sapiens |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. |
| REFERENCE | |
| AUTHORS | 1 Jiang,Y., Harlocker,S.L. and Secrist,H. |
| TITLE | Compositions and methods for the therapy and diagnosis of colon cancer |
| JOURNAL | Patent: WO 0196388-A 335 20-DEC-2001; |
| FEATURES | CORIXA CORPORATION (US) |
| source | Location/Qualifiers |
| | 1. .63 |
| | /organism="Homo sapiens" |
| | /db_xref="taxon:9606" |
| BASE COUNT | 9 a 2 c 4 g 44 t 4 others |
| ORIGIN | |
| Query Match | 0.6%; Score 30; DB 6; Length 63; |
| Best Local Similarity | 72.5%; Pred. No. 2.2e+04; |
| Matches | 37; Conservative 1; Mismatches 13; Indels 0; Gaps 0; |
| QY | 4475 TAAATTGAATGTAATTCATTAAAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4525 |
| Db | 11 TTTTGTAAAD 61 |
| RESULT 13 | |
| LOCUS | AF042300 73 bp DNA linear STS 26-JAN-1999 |
| DEFINITION | Homo sapiens chromosome 6 map 6p21.3, sequence tagged site. |
| ACCESSION | AF042300 |
| VERSION | AF042300.1 GI:4191287 |
| KEYWORDS | STS. |
| SOURCE | Homo sapiens. |
| ORGANISM | Homo sapiens |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. |
| REFERENCE | 1 (bases 1 to 73) |

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 20:46:07 ; Search time 659 Seconds
(Without alignments)
17841.704 Million cell updates/sec

Title: US-09-918-187-3
Perfect score: 5221
Sequence: 1 ataaaggggctgagga...aatctaaaaaaaaaaaaa 5221

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 2389434

Minimum DB seq length: 8
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

| | |
|-----|--|
| 1: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:* |
| 2: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:* |
| 3: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:* |
| 4: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:* |
| 5: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:* |
| 6: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:* |
| 7: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:* |
| 8: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:* |
| 9: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:* |
| 10: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:* |
| 11: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:* |
| 12: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:* |
| 13: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:* |
| 14: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:* |
| 15: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:* |
| 16: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:* |
| 17: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:* |
| 18: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:* |
| 19: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:* |
| 20: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:* |
| 21: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:* |
| 22: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:* |
| 23: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:* |
| 24: | /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| C 1 | 52.4 | 1.0 | 54 | 17 | AAT28167 |
| 2 | 49.4 | 0.9 | 51 | 22 | AAL28163 |
| 3 | 39 | 0.7 | 39 | 17 | AAT28105 |
| C 4 | 36.6 | 0.7 | 67 | 21 | AAAB7350 |
| 5 | 36.2 | 0.7 | 65 | 24 | ABN53617 |
| 6 | 34.2 | 0.7 | 97 | 16 | AAT21227 |
| C 7 | 33.8 | 0.6 | 83 | 24 | ABS23844 |
| 8 | 31 | 0.6 | 95 | 17 | AAT13984 |
| C 9 | 30.6 | 0.6 | 93 | 24 | ABN67931 |

| | | | | | | |
|------|------|-----|-----|----|----------|--------------------|
| C 10 | 30.4 | 0.6 | 64 | 22 | AAH48695 | Human G-protein su |
| 11 | 30 | 0.6 | 63 | 24 | ABL36746 | Human colon tumour |
| 12 | 30 | 0.6 | 65 | 24 | ABN28832 | Rat spliced transc |
| 13 | 29.8 | 0.6 | 68 | 21 | AAC17498 | Human secreted pro |
| 14 | 29.6 | 0.6 | 60 | 24 | ABN36661 | Human spliced tran |
| 15 | 29.4 | 0.6 | 73 | 24 | ABK45717 | CDNA encoding colo |
| 16 | 29.2 | 0.6 | 90 | 12 | AAQ11762 | Self-complementary |
| C 17 | 29 | 0.6 | 95 | 24 | ABS19692 | Human genome-deriv |
| C 18 | 28.8 | 0.6 | 50 | 18 | AAT65787 | Repeat sequence fr |
| C 19 | 28.8 | 0.6 | 65 | 18 | AAT65724 | Repeat sequence fr |
| C 20 | 28.8 | 0.6 | 65 | 18 | AAT66072 | (dc-da)n.(dc-dt)n |
| 21 | 28.8 | 0.6 | 75 | 21 | AAC29909 | Human secreted pro |
| C 22 | 28.2 | 0.5 | 51 | 18 | AAT65777 | Repeat sequence fr |
| C 23 | 28.2 | 0.5 | 86 | 22 | AAS23317 | Human prostate can |
| 24 | 28 | 0.5 | 62 | 16 | AAT20450 | Human gene signatu |
| C 25 | 27.8 | 0.5 | 57 | 18 | AAT65792 | Repeat sequence fr |
| C 26 | 27.6 | 0.5 | 73 | 13 | AAQ34140 | Sequence of a micr |
| C 27 | 27.6 | 0.5 | 74 | 24 | AAI69128 | Activated T-cell d |
| 28 | 27.6 | 0.5 | 87 | 21 | AAC18326 | Human secreted pro |
| 29 | 27.4 | 0.5 | 75 | 19 | AAV27232 | Hela gene sequence |
| C 30 | 27.4 | 0.5 | 75 | 22 | AAH43150 | SeqLPD#608pr51. S |
| C 31 | 27.2 | 0.5 | 50 | 18 | AAT65774 | Repeat sequence fr |
| C 32 | 27.2 | 0.5 | 55 | 24 | ABK53132 | HIV detection meth |
| C 33 | 27.2 | 0.5 | 55 | 24 | AAL45494 | HIV-1 RNA capture |
| C 34 | 27.2 | 0.5 | 90 | 22 | AAS23666 | Tetracycline promo |
| C 35 | 27 | 0.5 | 85 | 19 | AAV32418 | Homo sapiens clone |
| C 36 | 27 | 0.5 | 85 | 22 | AAF98434 | Human cDNA clone B |
| 37 | 27 | 0.5 | 87 | 24 | AAI69175 | Activated T-cell d |
| 38 | 27 | 0.5 | 99 | 21 | AAC11621 | Human secreted pro |
| C 39 | 27 | 0.5 | 100 | 15 | AAQ62573 | Candida albicans-s |
| C 40 | 26.8 | 0.5 | 46 | 18 | AAT65709 | Repeat sequence fr |
| C 41 | 26.8 | 0.5 | 49 | 18 | AAT65785 | Repeat sequence fr |
| C 42 | 26.8 | 0.5 | 53 | 18 | AAT65768 | Repeat sequence fr |
| C 43 | 26.8 | 0.5 | 54 | 13 | AAQ33870 | Microsatellite seq |
| C 44 | 26.8 | 0.5 | 56 | 18 | AAT66063 | (dc-da)n.(dc-dt)n |
| C 45 | 26.8 | 0.5 | 65 | 18 | AAT65730 | Repeat sequence fr |

ALIGNMENTS

| | |
|------------|---|
| RESULT 1 | |
| AAT28167/c | |
| ID | AAT28167 standard; DNA; 54 BP. |
| XX | |
| AC | AAT28167; |
| XX | |
| DT | 06-JAN-1997 (first entry) |
| XX | |
| DE | Senescence-related gene sequence 10F1. |
| XX | |
| KW | Human; fibroblast; AIDS; enhanced differential display; mRNA preparation; |
| KW | senescent cell; quiescent cell; dividing cell; senescence-related gene; |
| KW | gene expression; non-senescent cell; age-related lipofuscin; retina; EBD; |
| KW | therapy; liver spot; donor tissue; senescent melanocyte; melanin; |
| KW | hypopigmentation; ss. |
| XX | |
| OS | Synthetic. |
| XX | |
| PN | W09613610-A2. |
| XX | |
| PD | 09-MAY-1996. |
| XX | |
| PF | 24-AUG-1995; 95WO-US11230. |
| XX | |
| PR | 31-OCT-1994; 94US-0332420. |
| XX | |
| PA | (GERO-) GERON CORP. |
| XX | |
| PI | Feng J, Funk W, Hirsch KS, Linskens MHK, Villeponteau B; |
| PI | West MD; |
| XX | |
| DR | WPI; 1996-251464/25. |

XX Identifying, isolating and regulating senescence-related genes -
PT useful to ameliorate problems associated with accumulation of
PT senescent cells, e.g. age-related lipofuscin accumulation in the
PT retina and AIDS
XX
PS Claim 8; Page 51; 135pp; English.
XX
CC AAT28076-T28113, and AAT28131-T28173 represent novel senescent-related
CC gene sequences isolated from fibroblasts using the method of the
CC invention. In the method of the invention, mRNA is isolated from a
CC senescent cell, and a young quiescent cell, and the mRNAs are amplified
CC (using primers such as those shown in AAT28044-T28075) in separate
CC reaction mixtures. The amplified sequences are then separated by size or
CC charge, and the products are analysed to identify a gene from young
CC quiescent cells and dividing cells, that is present at a different level
CC from senescent cells. To enhance the method even more, it can be
CC performed in conjunction with an enhanced differential display (EDD)
CC method (an mRNA preparation method) on the fibroblasts. The method can
CC be used for the rapid and efficient identification and isolation of
CC senescence-related genes and gene products, and to detect and distinguish
CC between senescent and non-senescent cells. It can also be used to
CC destroy cells expressing senescence specific (or related) gene products,
CC and to screen for compounds capable of altering gene expression in
CC senescent cells. The method can also be used to ameliorate problems
CC associated with the accumulation of senescent cells such as age-related
CC lipofuscin accumulation in the retina, and in the treatment of AIDS.
CC Also, the method can be used to distinguish young cells from senescent
CC cells in donor tissue, which is useful in removing senescent melanocytes
CC overexpressing melanin which cause hypopigmentation, or liver spots.
XX
SO Sequence 54 BP; 17 A; 9 C; 9 G; 19 T; 0 other;
XX
Query Match 1.0%; Score 52.4; DB 17; Length 54;
Best Local Similarity 98.1%; Pred. No. 0.00039;
Matches 53; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1402 AGGATGCTAAAGATGATGATGTTACCCATTCAGTACAGTATCTTTAAAT 1455
DB 54 AGGATGCCAAGATGATGATGTTAACCCATTCAGTACAGTATCTTTAAAT 1
RESULT 2
AAL28163
ID AAL28163 standard; DNA; 51 BP.
XX
AC AAL28163;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #1371.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinasin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
XX
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.

1

XX
PI Shinkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 1772; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SO Sequence 51 BP; 11 A; 18 C; 15 G; 7 T; 0 other;
XX
Query Match 0.9%; Score 49.4; DB 22; Length 51;
Best Local Similarity 98.0%; Pred. No. 0.0025;
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 200 CCCCCTGGAAGTGAATCCCGCATCCGAGAGCAAGATGCCGCCACTTG 250
DB 1 CCCCCTGGAAGTGAATCCCGCATCCGAGAGCAAGATGCCGCCACTTG 51
RESULT 3
AAT28105
ID AAT28105 standard; DNA; 39 BP.
XX
AC AAT28105;
XX
DT 31-DEC-1996 (first entry)
XX
DE Probe 10F1 isolated from fibroblasts.
XX
KW Polymerase chain reaction; PCR; primer; amplify; human; fibroblast; AIDS;
KW enhanced differential display; EDD; mRNA preparation; senescent cell;
KW quiescent cell; dividing cell; senescence-related gene; gene expression;
KW non-senescent cell; age-related lipofuscin; retina; therapy; liver spot;
KW donor tissue; senescent melanocyte; melanin; hypopigmentation; ss.
XX
OS Synthetic.
XX
PN WO9613610-A2.
XX
PD 09-MAY-1996.
XX
PF 24-AUG-1995; 95WO-US11230.
XX
PR 31-OCT-1994; 94US-0332420.
XX
PA (GERO-) GERON CORP.
XX
PI Feng J, Funk W, Hirsch KS, Linskens MHK, Villeponteau B;
PI West MD;
XX
DR WPI; 1996-251464/25.
XX
PT Identifying, isolating and regulating senescence-related genes -

PT useful to ameliorate problems associated with accumulation of
PT senescent cells, e.g. age-related lipofuscin accumulation in the
PT retina and AIDS
XX
PS Claim 8; page 38; 135pp; English.
XX
CC AAT28076-T28113, and AAT28131-T28173 represent novel senescent-related
CC gene sequences isolated from fibroblasts using the method of the
CC invention. In the method of the invention, mRNA is isolated from a
CC senescent cell, and a young quiescent cell, and the mRNAs are amplified
CC (using primers such as those shown in AAT28044-T28075) in separate
CC reaction mixtures. The amplified sequences are then separated by size or
CC charge, and the products are analysed to identify a gene from young
CC quiescent cells and dividing cells, that is present at a different level
CC from senescent cells. To enhance the method even more, it can be
CC performed in conjunction with an enhanced differential display (EDD)
CC method (an mRNA preparation method) on the fibroblasts. The method can
CC be used for the rapid and efficient identification and isolation of
CC senescence-related genes and gene products, and to detect and distinguish
CC between senescent and non-senescent cells. It can also be used to
CC destroy cells expressing senescence specific (or related) gene products,
CC and to screen for compounds capable of altering gene expression in
CC senescent cells. The method can also be used to ameliorate problems
CC associated with the accumulation of senescent cells such as age-related
CC lipofuscin accumulation in the retina, and in the treatment of AIDS.
CC Also, the method can be used to distinguish young cells from senescent
CC cells in donor tissue, which is useful in removing senescent melanocytes
CC overexpressing melanin which cause hypopigmentation, or liver spots.
XX
SQ Sequence 39 BP; 11 A; 7 C; 6 G; 15 T; 0 other;

Query Match 0.7%; Score 39; DB 17; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1412 AGATGATGATGTTAAACCATTCACAGTACGATTTCTTTT 1450
 |||||
Db 1 AGATGATGATGTTAAACCATTCACAGTACGATTTCTTTT 39

RESULT 4
AAAB7350/C
ID AAAB7350 standard; DNA; 67 BP.
XX
AC AAAB7350;
XX
DT 08-JAN-2001 (first entry)
XX
DE Rat hepatocyte carcinogenesis biomarker nucleic acid SEQ ID NO:274.
XX
KW Rat; phenobarbital; carcinogenesis marker; carcinogenesis; detection;
KW identification; carcinogenic; probe; primer; ds.
XX
OS Rattus norvegicus.
XX
PN WO200044902-A2.
XX
PD 03-AUG-2000.
XX
PE 28-JAN-2000; 2000WO-US00503.
XX
PR 29-JAN-1999; 99US-0118078.
XX
PA (SEAR) SEARLE & CO G D.
XX
PI Bunch RT, Curtis SW, Rodi CP, Morris DL;
XX
DR WPI; 2000-505977/45.
XX
PT New nucleic acid encoding a carcinogenic biomarker, induced by
PT phenobarbital treatment of rat hepatocytes, useful for identifying
PT carcinogenic compounds -
XX

PS Claim 1; Page 137; 240pp; English.
XX
CC AAA87080 to AAA87656 represent nucleic acid sequences (N1) encoding a
CC carcinogenesis biomarkers. The carcinogenesis biomarkers are induced by
CC treating rat hepatocytes with phenobarbital. The nucleic acids are
CC useful for identifying carcinogenic compounds. The nucleic acid molecules
CC can be used to derive probes and/or primers for detecting or inducing
CC carcinogenesis, respectively.
XX
SQ Sequence 67 BP; 25 A; 11 C; 10 G; 21 T; 0 other;

Query Match 0.7%; Score 36.6; DB 21; Length 67;
Best Local Similarity 83.1%; Pred. No. 8.9;
Matches 54; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

OY 5136 TAATATCAACTCT--GAAGCTAATTGTACTAATCTGAGATGNGTTGTCAATAATAAA 5193
 |||||
Db 67 TAATATCAACTCTTGAGCCCTAATGTACTGATTGAGATTGCATTGTCCTAATAAAA 8

OY 5194 GTGAA 5198
 |||
Db 7 GTGGA 3

RESULT 5
ABN53617
ID ABN53617 standard; DNA; 65 BP.
XX
AC ABN53617;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:26365.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PE 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 26365; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini

CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 65 BP; 13 A; 22 C; 13 G; 17 T; 0 other;

Query Match 0.7%; Score 36.2; DB 24; Length 65;
Best Local Similarity 72.3%; Pred. No. 11;
Matches 47; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 521 TGCACGTTCTACACCTGGCTTGGGGGATCTCTACTATTGTGTCAGTCCCTGGGCATA 580
1 TGCACGCTCTACACTGCCCTCTCGGAGATTCTTCTACTACATGACCAGCGCTCGGCATC 60

OY 581 ACAGC 585
61 ACAGC 65

RESULT 6

AAT21227
ID AAT21227 standard; cDNA to mRNA; 97 BP.

AC AAT21227;
DT 01-AUG-1996 (first entry)

DE Human gene signature HUMGS02543.

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.

OS Homo sapiens.
PN WO9514772-A1.

PD 01-JUN-1995.

PF 11-NOV-1994; 94WO-JP01916.

PR 12-NOV-1993; 93JP-0355504.

PA (MATS/) MATSUBARA K.
(OKUB/) OKUBO K.

PI Matsuura K, Okubo K;

DR WPI; 1995-206931/27.

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 832; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-

CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX

SO Sequence 97 BP; 23 A; 18 C; 21 G; 35 T; 0 other;

Query Match 0.7%; Score 34.2; DB 16; Length 97;
Best Local Similarity 62.1%; Pred. No. 50;
Matches 54; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 4683 TGTGCACGACTGTGCTGTGACCCCTGGGCAAGTCACCTAATAAGTGCCCTCAGTT 4742
3 TCTGCTGCTTCCAGAGTGTGTGACCTTGGCTAAGTCACCTTAACCTTCTGATGTGATTT 62

OY 4743 TTGCTTCTGTTAAATGGGATTAATA 4769
63 CGCTTTTAAATAAGTGGCTGTGCTGA 89

RESULT 7

ABS23844/C
ID ABS23844 standard; DNA; 83 BP.

AC ABS23844;

DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe ORF from lung SEQ ID No 23835.

KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.

OS Homo sapiens.

PN WO200186003-A2.

PD 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US00665.

PR 04-FEB-2000; 2000US-180312P.

PR 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX Claim 4; SEQ ID No 23835; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived

PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR P-PSDB; ABP27300.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 7; Page 3532; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 93 BP; 44 A; 15 C; 12 G; 22 T; 0 other;

Query Match 0.6%; Score 30.6; DB 24; Length 93;
Best Local Similarity 62.3%; Pred. No. 4.7e+02;
Matches 48; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 5041 AGCATGCCCTTCTCTGATGGGTGGGATTTTCCCTTTTATGCGATATAGTACTT 5100
DB 77 ACGATGAATTTCTGTATTGTTGGTGGCTATTATTCTGTGTTGTTAGAAATTATTGAT 18
QY 5101 ACTGTGACAAAATAA 5117
DB 17 ACTAGTAGCAATACCAA 1

RESULT 10
AAH48695/c
ID AAH48695 standard; DNA; 64 BP.
XX
AC AAH48695;
XX
DT 19-OCT-2001 (first entry)
XX
DE Human G-protein subunit Gbeta3 Intron 10 DNA fragment #2.
XX
KW Human; G-protein subunit Gbeta3; variant; high blood pressure; asthma;
KW hypertension; cardiac infarction; coronary disease; heart disease;
KW circulation disease; diabetes; psychiatric disease; Crohn's disease;
KW immunological disease; psoriasis; colitis ulcerosa;
KW transplant rejection; ds.
XX
OS Homo sapiens.
XX
PN DE10030945-A1.
XX
PD 09-AUG-2001.
XX
PF 24-JUN-2000; 2000DE-1030945.
XX
PR 03-FEB-2000; 2000DE-1004681.
PR 21-FEB-2000; 2000DE-1007587.
XX
PA (SIFP/) SIFFERT W.

XX
PI Siffert W;
XX
DR WPI; 2001-490007/54.
XX
PT Use of sequence variants of the beta3 subunit of human G protein to
PT predict, detect, and determine drug targets for diseases including
PT heart, immunological and psychiatric diseases -
XX
PS Disclosure; Page 3; 14pp; German.
XX
CC This invention describes a novel use for sequence variants of the human
CC G-protein beta3 subunit to predict physiological and pathological
CC pathways in the human body. The invention may be used to detect, detect
CC predisposition for, give a prognosis of, or treat a variety of diseases
CC including high blood pressure, hypertension, cardiac infarction, coronary
CC disease and other heart or circulation diseases, diabetes, psychiatric
CC diseases, immunological diseases, asthma, psoriasis, Crohn's disease,
CC colitis ulcerosa, transplant rejection, HIV or Hepatitis B or C. This
CC sequence represents a fragment of the human G-protein Gbeta3 subunit
CC Intron 10 which is used to illustrate the method of the invention.
XX
SQ Sequence 64 BP; 24 A; 34 C; 1 G; 5 T; 0 other;

Query Match 0.6%; Score 30.4; DB 22; Length 64;
Best Local Similarity 85.0%; Pred. No. 4.2e+02;
Matches 34; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4076 AATGTATGTGTGGCTGTGGGTGGGTGGAGTGTGCTG 4115
DB 55 AGCTATATGTGTGGTGTGTGCTGTGGCTGTGTGTGTATG 16

RESULT 11
ABL36746
ID ABL36746 standard; cDNA; 63 BP.
XX
AC ABL36746;
XX
DT 08-APR-2002 (first entry)
XX
DE Human colon tumour antigen polynucleotide SEQ ID NO:335.
XX
KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200196388-A2.
XX
PD 20-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US18557.
XX
PR 09-JUN-2000; 2000US-210899P.
PR 20-FEB-2001; 2001US-270216P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Harlocker SL, Secrist H;
XX
DR WPI; 2002-114514/15.
XX
PT Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT useful for inhibiting development of cancer in patient -
XX
PS Claim 1; SEQ ID 335; 105pp; English.
XX
CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC which were isolated from human colon tumour and colon metastatic tumour
CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC production. (I) can be used for stimulating and/or expanding T cells

CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (1) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing
CC mutant species primers, or primers for use in genetic constructions. (1)
CC can be used in the diagnosis of a colon tumour.

XX
SQ Sequence 63 BP; 9 A; 2 C; 4 G; 44 T; 4 other;

Query Match 0.6%; Score 30; DB 24; Length 63;
Best Local Similarity 72.5%; Pred. No. 5.3e+02;
Matches 37; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

OY 4475 TAAATTTGAATGTAATTCGATTATTAAGTTTCTTTTCTTTTGGGTTAAAG 4525
1 ||| | | ||| ||| | | ||||| ||||| ||||| :
Db 11 TTNNTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTGTTAAAD 61

RESULT 12
ABN28832
ID ABN28832 standard; DNA; 65 BP.

XX
AC ABN28832;

XX
DT 15-JUL-2002 (first entry)

XX
DE Rat spliced transcript detection oligonucleotide SEQ ID NO:1580.

XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.

XX
OS Rattus norvegicus.

XX
PN WO200210449-A2.

XX
PD 07-FEB-2002.

XX
PF 20-JUL-2001; 2001WO-1B01903.

XX
PR 28-JUL-2000; 2000US-221607P.

XX
PR 02-MAY-2001; 2001US-287724P.

XX
PA (COMP-) COMPUGEN INC.

XX
VI Shoshan A, Wasserman A, Mintz E, Mintz L, Falgler S;

XX
DR WPI; 2002-257383/30.

XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -

XX
PS Example 1; SEQ ID 1580; 47pp; English.

XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent

CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 65 BP; 17 A; 23 C; 15 G; 10 T; 0 other;

Query Match 0.6%; Score 30; DB 24; Length 65;
Best Local Similarity 81.0%; Pred. No. 5.4e+02;
Matches 47; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

OY 201 CCCCTGGAAGTCATCCCGCATCCGAGCCCAAGATGCCCGCCACTTGTCTCAGCA 258
||| |||| |||| ||| ||||| ||||| ||||| ||| | ||||| ||
Db 6 CCCTGGAACTGACCCAGCATCCGA-CGCCAAGATGCCGCTCATACTGCAAGA 62

RESULT 13
AAC17498
ID AAC17498 standard; cDNA; 68 BP.

XX
AC AAC17498;

XX
DT 06-OCT-2000 (first entry)

XX
DE Human secreted protein 5' EST, SEQ ID NO: 21573.

XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX
OS Homo sapiens.

XX
PN EP1033401-A2.

XX
PD 06-SEP-2000.

XX
PF 21-FEB-2000; 2000EP-0200610.

XX
PR 26-FEB-1999; 99US-0122487.

XX
PA (GEST) GENSET.

XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX
DR WPI; 2000-500381/45.

XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 21573; 71pp + CD-ROM; English.

XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

XX
SQ Sequence 68 BP; 5 A; 9 C; 26 G; 28 T; 0 other;

Query Match 0.6%; Score 29.8; DB 21; Length 68;
Best Local Similarity 70.2%; Pred. No. 6.3e+02;
Matches 40; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 20:47:32 ; Search time 4414 Seconds
(without alignments)
19156.465 Million cell updates/sec

Title: US-09-918-187-3
Perfect score: 5221
Sequence: 1 ataaagggggctgaggaaa.....aatctaaaaaaaaaaaaaa 5221

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 357786

Minimum DB seq length: 8
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl1:*
10: gb_estl2:*
11: gb_htc:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 89.4 | 1.7 | 91 | 10 | BE546379 | BE546379 601071061 |
| 2 | 79.4 | 1.5 | 100 | 12 | BE695455 | BE695455 MRI-BT079 |
| 3 | 73.2 | 1.4 | 94 | 9 | AI339479 | AI339479 qt16d01.x |
| 4 | 73 | 1.4 | 73 | 12 | BG113100 | BG113100 602283756 |
| 5 | 66.4 | 1.3 | 68 | 9 | AA501922 | AA501922 ng54d08.s |
| 6 | 50 | 1.0 | 50 | 9 | AU103573 | AU103573 AU103573 |

| | | | | | | |
|----|------|-----|----|----|----------|--------------------|
| 7 | 50 | 1.0 | 50 | 9 | AU103578 | AU103578 AU103578 |
| 8 | 48.4 | 0.9 | 50 | 9 | AU103571 | AU103571 AU103571 |
| 9 | 48.4 | 0.9 | 50 | 9 | AU103576 | AU103576 AU103576 |
| 10 | 46.8 | 0.9 | 50 | 9 | AU103572 | AU103572 AU103572 |
| 11 | 45.8 | 0.9 | 58 | 9 | AA581009 | AA581009 nc83f06.r |
| 12 | 45.4 | 0.9 | 83 | 9 | AA108264 | AA108264 EST0007.r |
| 13 | 43.8 | 0.8 | 74 | 12 | BG255827 | BG255827 602368138 |
| 14 | 42 | 0.8 | 42 | 13 | BI223563 | BI223563 602941991 |
| 15 | 38.2 | 0.7 | 80 | 9 | AA466889 | AA466889 vf03h06.r |
| 16 | 36.2 | 0.7 | 96 | 9 | AI957503 | AI957503 cf92e11.x |
| 17 | 34.2 | 0.7 | 94 | 14 | C00785 | C00785 HDMS000254 |
| 18 | 33.8 | 0.6 | 89 | 9 | AI470686 | AI470686 t1j3f05.x |
| 19 | 32.8 | 0.6 | 69 | 14 | H83306 | H83306 ys89q09.r1 |
| 20 | 32.2 | 0.6 | 94 | 9 | AL634796 | AL634796 AL634796 |
| 21 | 32 | 0.6 | 32 | 13 | BI259338 | BI259338 602972590 |
| 22 | 32 | 0.6 | 69 | 9 | AJ500546 | AJ500546 AJ500546 |
| 23 | 31.6 | 0.6 | 98 | 13 | BI703413 | BI703413 fs90c05.y |
| 24 | 31.4 | 0.6 | 88 | 9 | AI545503 | AI545503 fb66c12.x |
| 25 | 31.4 | 0.6 | 93 | 9 | AA928844 | AA928844 on99f12.s |
| 26 | 31.2 | 0.6 | 60 | 9 | AI523624 | AI523624 t995b12.x |
| 27 | 31.2 | 0.6 | 76 | 9 | AI583160 | AI583160 tr98d01.x |
| 28 | 31.2 | 0.6 | 85 | 9 | AA808096 | AA808096 oc39c06.s |
| 29 | 30.8 | 0.6 | 75 | 17 | CNS015N0 | AL105654 Drosoph11 |
| 30 | 30.8 | 0.6 | 79 | 14 | BQ267226 | BQ267226 1j99g02.y |
| 31 | 30.8 | 0.6 | 83 | 9 | AI061323 | AI061323 an32h12.x |
| 32 | 30.8 | 0.6 | 84 | 17 | CNS032I6 | AL267639 Telraodon |
| 33 | 30.6 | 0.6 | 79 | 10 | AW085980 | AW085980 xc76e02.x |
| 34 | 30.6 | 0.6 | 83 | 13 | BI322346 | BI322346 kx19h06.y |
| 35 | 30.4 | 0.6 | 52 | 9 | AA782923 | AA782923 a162c05.s |
| 36 | 30.4 | 0.6 | 85 | 10 | AW263326 | AW263326 xq86g03.x |
| 37 | 30.4 | 0.6 | 94 | 13 | BI493030 | BI493030 df32e06.w |
| 38 | 30.4 | 0.6 | 98 | 9 | AA472898 | AA472898 vb30h11.r |
| 39 | 30.2 | 0.6 | 60 | 14 | R85541 | R85541 y038c03.s1 |
| 40 | 30.2 | 0.6 | 87 | 9 | AI539690 | AI539690 tp71h01.x |
| 41 | 30.2 | 0.6 | 87 | 14 | H95782 | H95782 yv18c06.s1 |
| 42 | 30.2 | 0.6 | 95 | 9 | AU053988 | AU053988 AU053988 |
| 43 | 30.2 | 0.6 | 96 | 9 | AI436732 | AI436732 th77a10.x |
| 44 | 30 | 0.6 | 59 | 14 | N24179 | N24179 yx69c08.s1 |
| 45 | 30 | 0.6 | 93 | 9 | AI429297 | AI429297 mm44g08.x |

ALIGNMENTS

RESULT 1
BE546379
LOCUS
DEFINITION 601071061F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3457239 5', mRNA sequence.
ACCESSION BE546379
VERSION BE546379.1 GI:9775024
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 91)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaphs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: LHAM8446 row: 1 column: 16
High quality sequence stop: 91.
Location/Qualifiers
1..91

FEATURES
source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3457239"
/clone_1lb="NH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/Note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
BASE COUNT      22 a      24 c      26 g      19 t
ORIGIN

Query Match
Best Local Similarity  98.9%; Score 89.4; DB 10; Length 91;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  3227 CAGCTTGAGCCAGTGGCCAGCCACAGAGCAAGAGGGTTATTTCAGTCCCTCTCT 3286
      |||||||
Db    1 CACGTTGAGCCAGTGGCCAGCCACAGAGCAAGAGGGTTATTTCAGTCCCTCTCT 60

QY  3287 CTGGGTCAGAACAGAGGCGCATGCTGAATGC 3317
      |||||||
Db    61 CTGGGTCAGAACAGAGGCGCATGCTGAATGC 91

RESULT 2
BE695455      100 bp      mRNA      linear      EST 11-SEP-2000
LOCUS
DEFINITION   MR1-BT0796-160600-001-c10 BT0796 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE695455
VERSION     BE695455.1 GI:10082615
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
JOURNAL
MEDLINE
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-MR1-BT0796-160
            600-001-c10&tl3=2000-06-16&tl4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 13
            High quality sequence stop: 100.
            Location/Qualifiers
                1..100
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_1lb="BT0796"
                /dev_stage="Adult"
                /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                SmaI; A mini-library was made by cloning products derived
                from ORFESTES PCR (U.S. Letters Patent application No. 196
                ,716 - Ludwig Institute for Cancer Research) profiles
```

```
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      24 a      30 c      23 g      23 t
ORIGIN

Query Match
Best Local Similarity  97.8%; Score 79.4; DB 12; Length 100;
Matches 91; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY  1191 TCACCACATCTTCATTCATTCATTCGCGCCCGCTCGCTGCGCTATGACCGGAAGAAAG 1250
      |||||||
Db    9 TTACCACATCTTCATTCATTCATTCGCGCCCGCTCGCTGCGCTATGACCGGAAGAAAG 67

QY  1251 TCTCCAGGCGCCGACATCTGGCCAGATTAATA 1283
      |||||||
Db    68 TCTCCAGGCGCCGACATCTGGCCAGATTAATA 100
```

```
RESULT 3
AI339479/c
LOCUS
DEFINITION   qtl6d01.x1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1947745 3'
similar to TR:000767 000767 ACYL-COA DESATURASE 1 ; , mRNA sequence.
AI339479
ACCESSION   AI339479.1 GI:4076406
VERSION     AI339479.1 GI:4076406
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS     1 (bases 1 to 94)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgaaps-remail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
JOURNAL
COMMENT     Trace considered overall poor quality
            Seq primer: -40up from Gibco
            High quality sequence stop: 1.
            Location/Qualifiers
                1..94
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:1947745"
                /clone_1lb="NCI_CGAP_GC4"
                /tissue_type="pooled germ cell tumors"
                /lab_host="DH10B"
                /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                polylinker; 1st strand cDNA was prepared from 3 pooled
                germ cell tumors, and was then primed with a Not I -
                oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                RI adaptors (Pharmacia), digested with Not I and cloned
                into the Not I and Eco RI sites of the modified pT7T3
                vector. Library is normalized. Library was constructed by
                Bento Soares and M. Fatima Bernaldo."
```

```
BASE COUNT      26 a      25 c      24 g      19 t
ORIGIN

Query Match
Best Local Similarity  86.2%; Score 73.2; DB 9; Length 94;
Matches 81; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

OY 992 CGATATGCTGTGCTTATATGCCACCTGGCTGTGACAGTGTGCCCCACCTCTTCGA 1051
 DB 94 CGATATGCTGTGCTTATATCTCGCTGGGTGTAAACAGTGTAGCGCCCTCATCTGA 35
 OY 1052 TATCGTCCATTAGACAGACATTTAGCCCCGGG 1085
 DB 34 TATCGTCCATTAGACAGACATTTAGCCTCCGGG 1

RESULT 4
 LOCUS BG113100 73 bp mRNA linear EST 30-JAN-2001
 DEFINITION 602283756F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4371325 5',
 mRNA sequence.
 ACCESSION BG113100
 VERSION BG113100.1 GI:12606606
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 73)
 NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM10029 row: P column: 14
 High quality sequence stop: 73.
 Location/Qualifiers
 1. 73
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4371325"
 /clone_1lb="NIH_MGC_86"
 /tissue_type="osteosarcoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: bone; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.533 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 12 a 20 c 19 g 22 t
 ORIGIN
 Query Match 1.4%; Score 73; DB 12; Length 73;
 Best Local Similarity 100.0%; Pred. No. 0.00026;
 Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3263 GGGTTTATTTTCAGTCCCTCTCTGGGTCAGAACCCAGAGGGCATGCTGAATGCCCTT 3322
 DB 1 GGGTTTATTTTCAGTCCCTCTCTGGGTCAGAACCCAGAGGGCATGCTGAATGCCCTT 60
 OY 3323 GCTTACTTGGTGA 3335
 DB 61 GCTTACTTGGTGA 73

RESULT 5
 LOCUS AA501922 68 bp mRNA linear EST 18-AUG-1997
 DEFINITION ng54d08.s1 NCI_CGAP_L12 Homo sapiens cDNA clone IMAGE:938607, mRNA
 sequence.
 ACCESSION AA501922
 VERSION AA501922.1 GI:2236889
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 68)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: David E. Kleiner, M.D., Ph.D., Rodrigo F.
 Chuagui, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: David B. Krizman, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/dbp/image/image.html
 Insert Length: 607 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 51.
 Location/Qualifiers
 1. 68
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:938607"
 /clone_1lb="NCI_CGAP_L12"
 /sex="male"
 /tissue_type="liver"
 /lab_host="DH10B"
 /note="Vector: pAMP10; mRNA made from invasive
 hepatocellular carcinoma, cDNA made by oligo-dT priming.
 Non-directionally cloned. Size-selected on agarose gel,
 average insert size 600 bp. Reference: Krizman et al.
 (1996) Cancer Research 56:5380-5383."

BASE COUNT 12 a 20 c 17 g 19 t
 ORIGIN

Query Match 1.3%; Score 66.4; DB 9; Length 68;
 Best Local Similarity 98.5%; Pred. No. 0.0059;
 Matches 67; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 2944 GCAGCTCTTTGTGTGTTATTCAGAGGACAGTGTGCTGTCCAGGACGCTCCCTCC 3003
 DB 1 GCAGCTCTTTGTGTGTTATTCAGAGGACAGTGTGCTGTCCAGGACGCTCCCTCC 60
 OY 3004 TGCACACA 3011
 DB 61 TGCACACA 68

RESULT 6
 LOCUS AU103573 50 bp mRNA linear EST 30-AUG-2001
 DEFINITION AU103573 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 HEP13102, mRNA sequence.
 ACCESSION AU103573
 VERSION AU103573.1 GI:13553094
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 50)
 Suzuki, Y., Talra, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
 'H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki
 'Y., Nakamura, Y., Suyama, A. and Sugano, S.
 Diverse transcriptional initiation revealed by fine, large-scale
 mapping of mRNA start sites
 EMBO Rep. 2 (5), 388-393 (2001)
 JOURNAL 21270072
 MEDLINE
 COMMENT Contact: Yutaka Suzuki

Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: y Suzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
, S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP13102"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and
dimethylfumarate treated U937 cells"

BASE COUNT
8 a 18 c 12 g 12 t

ORIGIN
Query Match 1.0%; Score 50; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 ACCGTCACCCGCTGCGACGCTGCTTAATCCCGCTCGGGACCT 78
|||||
DB 1 ACGGTCACCCGCTGCGACGCTGCTTAATCCCGCTCGGGACCT 50

RESULT 7
AUI03578 50 bp mRNA linear EST 30-AUG-2001
LOCUS AUI03578 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION ADSE01417, mRNA sequence.
ACCESSION AUI03578
VERSION AUI03578.1 GI:13553099
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 50)
Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki
, Y., Nakamura, Y., Suyama, A. and Sugano, S.
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)
21270072
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: y Suzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
, S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ADSE01417"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and
dimethylfumarate treated U937 cells"

BASE COUNT
10 a 19 c 17 g 4 t

ORIGIN
Query Match 1.0%; Score 50; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 ACCGGCGCTAGCGCGACACGCTAGCGTGCAAGCGCGCGCTCAG 134
|||||
DB 1 ACGCGGCTAGCGCGACACGCTAGCGTGCAAGCGCGCGCTCAG 50

RESULT 8
AUI03571 50 bp mRNA linear EST 30-AUG-2001
LOCUS AUI03571 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION AUI03571 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP07904, mRNA sequence.
ACCESSION AUI03571
VERSION AUI03571
KEYWORDS EST.
SOURCE AUI03571.1 GI:13553092
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 50)
Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki
, Y., Nakamura, Y., Suyama, A. and Sugano, S.
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)
21270072
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: y Suzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
, S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP07904"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and
dimethylfumarate treated U937 cells"

BASE COUNT
9 a 19 c 17 g 5 t

ORIGIN
Query Match 0.9%; Score 48.4; DB 9; Length 50;
Best Local Similarity 98.0%; Pred. No. 32;
Matches 49; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 83 GCACCGCGCTAGCGCGACACGCTAGCGTGCAAGCGCGCGCTC 132
|||||
DB 1 GCACCGCGCTAGCGCGACACGCTAGCGTGCAAGCGCGCGCTC 50

RESULT 9
AUI03576 50 bp mRNA linear EST 30-AUG-2001
LOCUS AUI03576 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION AUI03576 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP21416, mRNA sequence.
ACCESSION AUI03576
VERSION AUI03576.1 GI:13553097
KEYWORDS EST.
SOURCE AUI03576.1 GI:13553097
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 50)
Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki
, Y., Nakamura, Y., Suyama, A. and Sugano, S.
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)
21270072
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yasukie@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
, S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP21416"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and
dimethylfumarate treated U937 cells"

BASE COUNT
8 a 19 c 14 g 9 t

ORIGIN

Query Match 0.9%; Score 48.4; DB 9; Length 50;
Best Local Similarity 98.0%; Pred. No. 32;
Matches 49; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 51 AGCCTTAATTCCTCGGCTCGGACCTCCACGACCGCGCTAGCGCG 100
|||||
Db 1 AGCCTTAATTCCTCGGCTCGGACCTTCACGACCGCGCTAGCGCG 50

RESULT 10
AU103572 50 bp mRNA linear EST 30-AUG-2001

LOCUS AU103572 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION HEP12765, mRNA sequence.
ACCESSION AU103572
VERSION AU103572.1 GI:13553093
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 50)
Suzuki, Y., Talra, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki
, Y., Nakamura, Y., Suyama, A. and Sugano, S.
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)

TITLE
JOURNAL
MEDLINE
COMMENT
CONTACT: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yasukie@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
, S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP12765"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and
dimethylfumarate treated U937 cells"

BASE COUNT
8 a 16 c 12 g 14 t

ORIGIN

Query Match 0.9%; Score 46.8; DB 9; Length 50;
Best Local Similarity 96.0%; Pred. No. 67;
Matches 48; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 29 ACGTCACCGCTTGCAGCTTAGCCTTAATTCCTCGGCTCGGAGACT 78
|||||
Db 1 ACGTCACCGCTTGCAGCTTAGCCTTAATTCCTCGGCTCGGAGACT 50

RESULT 11

AA581009 58 bp mRNA linear EST 05-JAN-1998
nc83f06.r1 NCI_CGAP_GC1 Homo sapiens cDNA clone IMAGE:797411 5',
similar to SW:AC02_MOUSE P13011 ACYL-COA DESATURASE 2 ;, mRNA
sequence.

ACCESSION AA581009
VERSION AA581009.1 GI:2358781
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 58)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies Inc., David Krizman,
Ph.D.

FEATURES

CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html
Seq primer: -28ml3 rev1 EF from Amersham
High quality sequence stop: 1.
location/Qualifiers

FEATURES

source
1..58
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:797411"
/clone_lib="NCI_CGAP_GC1"
/tissue_type="bulk germ cell seminoma"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT2; Site_1: SalI; Site_2: NotI;
Cloned unidirectionally. Primer: Oligo dt."

BASE COUNT
9 a 14 c 21 g 14 t

ORIGIN

Query Match 0.9%; Score 45.8; DB 9; Length 58;
Best Local Similarity 87.7%; Pred. No. 1e+02;
Matches 50; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1003 GGTGCTTAATGACCTGCTGGTGAACAGTCTGCCACCTTCGAGATCGTCC 1059
|||||
Db 1 GGTGCTTAATGACCTGCTGGTGAACAGTCTGCCACCTTCGAGATCGTCC 57

FEATURES

AA108264 83 bp mRNA linear EST 05-SEP-1997
EST0007 rat lambda ZAPIT library (C.P.Hamel) Rattus norvegicus cDNA
clone PC038 5' similar to Stearyl-CoA denaturase, mRNA sequence.
ACCESSION AA108264
VERSION AA108264.1 GI:1659732
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 83)
Soto-Prior, A., Lavigne-Rebillard, M., Lenoir, M., Ripoll, C.,
Rebillard, G., Vago, P., Pujol, R. and Hamel, C.P.
Identification of preferentially expressed cochlear genes by
systematic sequencing of a rat cochlea cDNA library

FEATURES

JOURNAL
MEDLINE
97364947

COMMENT Contact: Hamel, CP
Laboratoire de Neurobiologie de l'Audition - INSERM U254
Institut National de la Sante et de la Recherche Medicale
C.H.R. St-Charles, 34295 Montpellier cedex, France
Tel: (33) 04 67 33 69 75
Fax: (33) 04 67 52 56 01
Email: biomol@met.fr
PCR Primers
FORWARD: pUC/M13 reverse primer
BACKWARD: pUC/M13 forward primer
Seq primer: pUC/M13 reverse primer
High quality sequence stop: 83.
Location/Qualifiers
1..83
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/clone="pCO38"
/clone_1lb="rat lambda ZAPII library (C.P.Hamel)"
/dev_stage="postnatal day 24"
/lab_host="Escherichia coli"
/note="Organ: cochlea; Vector: lambda ZAPII; Site_1: EcoRI
; Site_2: XhoI"

BASE COUNT 13 a 34 c 17 g 18 t 1 others

ORIGIN

Query Match 0.9%; Score 45.4; DB 9; Length 83;
Best Local Similarity 89.1%; Pred. No. 1.1e+02;
Matches 49; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 626 CGGCTGCCCTACGGCTCTTCATCATTCGCCACACATGCGCATTCAGATG 680
|||||
DB 4 CGGCTGCCCTACGGCTCTTCATCATTCGCCACACACATGCGCATTCAGACG 58

RESULT 13
BG255827 74 bp mRNA linear EST 13-FEB-2001
LOCUS 60236813BF1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4476417 5',
DEFINITION mRNA sequence.
ACCESSION BG255827
VERSION BG255827.1 GI:12765643
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 74)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM10303 row: k column: 10
High quality sequence stop: 41.
Location/Qualifiers
1..74
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4476417"
/clone_1lb="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for

FEATURES
source

full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 11 a 14 c 24 g 25 t

ORIGIN

Query Match 0.8%; Score 43.8; DB 12; Length 74;
Best Local Similarity 95.7%; Pred. No. 2.5e+02;
Matches 45; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2283 CCATCTTCAGGATATGCTCTTCCCTCATATGTAATAGATGGCTG 2329
|||||
DB 1 CCATCTTCAGGATATGCTCTTCCCTCATATGTAATAGATGGCTG 47

RESULT 14
BI223563 42 bp mRNA linear EST 11-JUL-2001
LOCUS 602941991F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5104919 5',
DEFINITION mRNA sequence.
ACCESSION BI223563
VERSION BI223563.1 GI:14677007
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 42)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM11253 row: f column: 24
High quality sequence stop: 42.
Location/Qualifiers
1..42
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5104919"
/clone_1lb="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT 10 a 11 c 12 g 9 t

ORIGIN

Query Match 0.8%; Score 42; DB 13; Length 42;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3294 AGAACGAGAGGGCATGCTGATGCCCTGCTTACTTGCTGA 3335
|||||
DB 1 AGAACGAGAGGGCATGCTGATGCCCTGCTTACTTGCTGA 42

RESULT 15
AA466889 80 bp mRNA linear EST 11-JUN-1997
LOCUS AA466889
DEFINITION v03h06.r1 Knowles Solter mouse blastocyst B3 Mus musculus CDNA
clone IMAGE:834683 5' similar to gb:M21285 Mouse steryl-1-CoA
desaturase gene (MOUSE);, mRNA sequence.
ACCESSION AA466889
VERSION AA466889.1 GI:2193029

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 20:58:57 ; Search time 116 Seconds

(Without alignments)
13803.098 Million cell updates/sec

Title: US-09-918-187-3

Perfect score: 5221

Sequence: 1 atataaaggggctgaggaataa.....aatctaaaaaaaaaaaaa 5221

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 684418

Minimum DB seq length: 8

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------------------|--------------------|
| 1 | 39 | 0.7 | 39 | 1 | US-08-332-420-62 | Sequence 62, Appl |
| 2 | 28.8 | 0.6 | 50 | 1 | US-08-222-177A-367 | Sequence 367, App |
| 3 | 28.8 | 0.6 | 65 | 1 | US-08-222-177A-116 | Sequence 116, App |
| 4 | 28.8 | 0.6 | 65 | 1 | US-08-222-177A-421 | Sequence 421, App |
| 5 | 28.2 | 0.5 | 51 | 1 | US-08-222-177A-337 | Sequence 337, App |
| 6 | 27.8 | 0.5 | 57 | 1 | US-08-222-177A-382 | Sequence 382, App |
| 7 | 27.4 | 0.5 | 75 | 3 | US-09-284-782-20 | Sequence 20, Appl |
| 8 | 27.2 | 0.5 | 50 | 1 | US-08-222-177A-328 | Sequence 328, App |
| 9 | 27 | 0.5 | 100 | 1 | US-08-145-705A-17 | Sequence 17, Appl |
| 10 | 26.8 | 0.5 | 46 | 1 | US-08-222-177A-71 | Sequence 71, Appl |
| 11 | 26.8 | 0.5 | 49 | 1 | US-08-222-177A-361 | Sequence 361, App |
| 12 | 26.8 | 0.5 | 53 | 1 | US-08-222-177A-304 | Sequence 304, App |
| 13 | 26.8 | 0.5 | 56 | 1 | US-08-222-177A-412 | Sequence 412, App |
| 14 | 26.8 | 0.5 | 65 | 1 | US-08-222-177A-134 | Sequence 134, App |
| 15 | 26.8 | 0.5 | 91 | 1 | US-08-222-177A-107 | Sequence 107, App |
| 16 | 26.6 | 0.5 | 89 | 1 | US-08-222-177A-98 | Sequence 98, Appl |
| 17 | 26.4 | 0.5 | 56 | 1 | US-08-222-177A-394 | Sequence 394, App |
| 18 | 26.4 | 0.5 | 62 | 6 | 5506118-3 | Patent No. 5506118 |
| 19 | 26.4 | 0.5 | 69 | 4 | US-09-269-911A-8 | Sequence 8, Appli |
| 20 | 26.4 | 0.5 | 84 | 3 | US-09-284-782-26 | Sequence 26, Appl |
| 21 | 26.2 | 0.5 | 51 | 1 | US-08-222-177A-89 | Sequence 89, Appl |
| 22 | 26.2 | 0.5 | 51 | 1 | US-08-222-177A-317 | Sequence 317, App |
| 23 | 26 | 0.5 | 83 | 1 | US-08-120-827-100 | Sequence 100, App |
| 24 | 26 | 0.5 | 83 | 1 | US-08-478-675-100 | Sequence 100, App |
| 25 | 26 | 0.5 | 92 | 5 | PCT-US93-06251-39 | Sequence 39, Appl |
| 26 | 25.8 | 0.5 | 72 | 1 | US-08-222-177A-131 | Sequence 131, App |
| 27 | 25.8 | 0.5 | 72 | 1 | US-08-222-177A-427 | Sequence 427, App |

| | | | | | | |
|----|------|-----|----|---|--------------------|-------------------|
| 28 | 25.8 | 0.5 | 95 | 1 | US-08-518-878B-3 | Sequence 3, Appl1 |
| 29 | 25.8 | 0.5 | 95 | 1 | US-08-294-522B-3 | Sequence 3, Appl1 |
| 30 | 25.8 | 0.5 | 95 | 2 | US-08-807-861A-3 | Sequence 3, Appl1 |
| 31 | 25.8 | 0.5 | 95 | 2 | US-08-470-868A-3 | Sequence 3, Appl1 |
| 32 | 25.8 | 0.5 | 95 | 3 | US-09-210-681-3 | Sequence 3, Appl1 |
| 33 | 25.6 | 0.5 | 95 | 3 | US-08-946-719A-3 | Sequence 3, Appl1 |
| 34 | 25.6 | 0.5 | 51 | 1 | US-08-222-177A-391 | Sequence 391, App |
| 35 | 25.6 | 0.5 | 62 | 1 | US-08-222-177A-56 | Sequence 56, Appl |
| 36 | 25.6 | 0.5 | 80 | 4 | US-09-284-627-15 | Sequence 15, Appl |
| 37 | 25.4 | 0.5 | 40 | 1 | US-08-222-177A-175 | Sequence 175, App |
| 38 | 25.4 | 0.5 | 45 | 1 | US-08-222-177A-157 | Sequence 157, App |
| 39 | 25.4 | 0.5 | 54 | 1 | US-08-469-802B-27 | Sequence 27, Appl |
| 40 | 25.4 | 0.5 | 54 | 2 | US-08-267-803B-45 | Sequence 45, Appl |
| 41 | 25.4 | 0.5 | 75 | 2 | US-08-776-944-13 | Sequence 13, Appl |
| 42 | 25.2 | 0.5 | 38 | 1 | US-08-222-177A-198 | Sequence 198, App |
| 43 | 25.2 | 0.5 | 39 | 1 | US-08-222-177A-137 | Sequence 137, App |
| 44 | 25.2 | 0.5 | 40 | 1 | US-08-222-177A-119 | Sequence 119, App |
| 45 | 25.2 | 0.5 | 40 | 1 | US-08-222-177A-400 | Sequence 400, App |

ALIGNMENTS

RESULT 1
US-08-332-420-62
; Sequence 62, Application US/08332420
; Patent No. 5744300
; GENERAL INFORMATION:
; APPLICANT: Maarten H.K. Linskens, et al.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR THE
; TITLE OF INVENTION: IDENTIFICATION AND REGULATION
; TITLE OF INVENTION: OF SENESENCE-RELATED GENES
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,420
; FILING DATE: October 31, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/235,180
; FILING DATE: April 29, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 210/041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-332-420-62

Query Match 0.7%; Score 39; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1412 AGATGATGATGTTAACCCATTCCAGTACAGTATCTTTT 1450

1 APPLICANT: Elmat, Paz
 2 TITLE OF INVENTION: GENE IDENTIFICATION METHOD
 3 NUMBER OF SEQUENCES: 36
 4 CORRESPONDENCE ADDRESS:
 5 ADDRESSEE: Kohn & Associates
 6 STREET: 30500 No. 605711thwestern Highway, Suite 410
 7 CITY: Farmington Hills
 8 STATE: Michigan
 9 COUNTRY: US
 10 ZIP: 48334
 11
 12 COMPUTER READABLE FORM:
 13 MEDIUM TYPE: Floppy disk
 14 COMPUTER: IBM PC compatible
 15 OPERATING SYSTEM: PC-DOS/MS-DOS
 16 SOFTWARE: PatentIn Release #1.0, Version #1.30
 17 CURRENT APPLICATION DATA:
 18 APPLICATION NUMBER: US/09/284,782
 19 FILING DATE:
 20
 21 CLASSIFICATION: 435
 22 ATTORNEY/AGENT INFORMATION:
 23 NAME: Montgomery, Ilene N.
 24 REGISTRATION NUMBER: 38,972
 25 REFERENCE/DOCKET NUMBER: 0168-00022
 26 TELECOMMUNICATION INFORMATION:
 27 TELEPHONE: (248) 539-5050
 28 TELEFAX: (248) 539-5055
 29 INFORMATION FOR SEQ ID NO: 20:
 30 SEQUENCE CHARACTERISTICS:
 31 LENGTH: 75 base pairs
 32 TYPE: nucleic acid
 33 STRANDEDNESS: single
 34 TOPOLOGY: linear
 35 MOLECULE TYPE: CDNA
 36 US-09-284-782-20

```

Query Match          0.5%; Score 27.4; DB 3; Length 75;
Best Local Similarity 65.6%; Pred. No. 2.6e+02;
Matches 40; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 4680 CACTGTCCACTGACTTGTGCTGTGTGACCCCTGGGCAAGTCACTTAAGGTCCTCA 4739
      1 111 111 1111 11111 1 1 11 1111111 1 1 111111
DB 9 CTCTGGGCTCTTGACCAAGCTGTGGCGGCATGGGCTAAGTCACTTGCCCTTGGAGCCTCT 68

QY 4740 G 4740
      1
DB 69 G 69

RESULT 8
US-08-222-177A-328/C
; Sequence 328, Application US/08222177A
; Patent No. 5582579
;
; GENERAL INFORMATION:
; APPLICANT: Weber, James L.
; TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
; TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 460
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Demilt Ross & Stevens, S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53717-1914
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,177A
; FILING DATE:
;
; CLASSIFICATION: 435

```

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER:  US 07/341,562
3      FILING DATE:  21-APR-1989
4      ATTORNEY/AGENT INFORMATION:
5      NAME:  Sara, Charles S.
6      REGISTRATION NUMBER:  30,492
7      REFERENCE/DOCKET NUMBER:  09865.601
8      TELECOMMUNICATION INFORMATION:
9      TELEPHONE:  (608) 831-2100
10     TELEFAX:  (608) 831-2106
11     TELEX:
12     INFORMATION FOR SEQ ID NO:  328:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH:  50 base pairs
15     TYPE:  nucleic acid
16     STRANDEDNESS:  double
17     TOPOLOGY:  linear
18     MOLECULE TYPE:  DNA (genomic)
19     IMMEDIATE SOURCE:
20     CLONE:  mfd103rs
21     US-08-222-177A-328

```

| | Query Match | Similarity | Score | DB 1; | length |
|------------|---|--|--------------------|-------|--------|
| Best Local | 32; | 80.0%; | Pred. No. 2.3e+02; | | |
| Matches | Conservative | 0; | Mismatches | 8; | Indels |
| | | | | | Gaps |
| QY | 4078 TGTATGTCCTGTTGGTGTGCCTCCAGTGTCGTCT | 4117 | | | |
| DB | 48 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 9 | | | |

```

1  RESULT 9
2  US-08-145-705A-17/c
3  ; Sequence 17, Application US/08145705A
4  ; Patent No. 5489513
5  ;
6  ; GENERAL INFORMATION:
7  ; APPLICANT: Springer, Wolfgang; Plempel, Manfred;
8  ; APPLICANT: L bberding, Antonius
9  ; TITLE OF INVENTION: SPECIFIC GENE PROBES AND
10 ; TITLE OF INVENTION: PROCESSES FOR THE DIAGNOSTIC
11 ; TITLE OF INVENTION: INVESTIGATION OF CANDIDA
12 ; TITLE OF INVENTION: ALBICANS
13 ; NUMBER OF SEQUENCES: 44
14 ;
15 ; CORRESPONDENCE ADDRESS:
16 ; ADDRESSEE: SPRUNG HORN KRAMER & WOODS
17 ; STREET: 660 White Plains Road
18 ; CITY: Tarrytown
19 ; STATE: New York
20 ; COUNTRY: U.S.A.
21 ; ZIP: 10591-5144
22 ;
23 ; COMPUTER READABLE FORM:
24 ; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 MB storage
25 ; COMPUTER: NEC Powermate 1 plus
26 ; OPERATING SYSTEM: DOS
27 ; SOFTWARE: Wordperfect 5.1
28 ;
29 ; CURRENT APPLICATION DATA:
30 ; APPLICATION NUMBER: US/08/145,705A
31 ; FILING DATE: October 28, 1993
32 ; CLASSIFICATION: 536
33 ;
34 ; PRIOR APPLICATION DATA:
35 ; APPLICATION NUMBER: German P 42 36 708.5
36 ; FILING DATE: October 30, 1992
37 ; ATTORNEY/AGENT INFORMATION:
38 ; NAME: Kurt G. Biscoe
39 ; REGISTRATION NUMBER: 33,141
40 ; REFERENCE/DOCKET NUMBER: Bayer 8885-KGB
41 ; TELECOMMUNICATION INFORMATION:
42 ; TELEPHONE: (914) 332-1700
43 ; TELEFAX: (914) 332-1844
44 ;
45 ; TELEX:
46 ; INFORMATION FOR SEQ ID NO: 17:
47 ; SEQUENCE CHARACTERISTICS:
48 ; LENGTH: 100 base pairs
49 ;
50 ;

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Candida albicans
US-08-145-705A-17

Query Match 0.5%; Score 27; DB 1; Length 100;
Best Local Similarity 62.7%; Pred. No. 4.1e+02;
Matches 42; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 2628 GTTACTGAGTCACTTATTCGAGCTCTTAATAAATAACTAGATATTAGTCCATTCAT 2687
1 111 1 111 1 111 1 111 1 1 111 1 11 11
Db 91 GATACCAAGACATTAATGGTAGTCTCCAACATAATCAATAAATTATCTATATTAGT 32

QY 2688 AATTAGT 2694
1 1111
Db 31 AGTTAGT 25

RESULT 10

US-08-222-177A-71/c
Sequence 71, Application US/08222177A
Patent No. 5582979

GENERAL INFORMATION:

APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:

INFORMATION FOR SEQ ID NO: 71:

SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-222-177A-71

Query Match 0.5%; Score 26.8; DB 1; Length 46;
Best Local Similarity 81.6%; Pred. No. 2.8e+02;
Matches 31; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4078 TGTATGTGTGGCTGTGGTGTGGGTGGAGTGTGTCTG 4115

Db 44 TGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 7
1111111111 1111 1111 1111 11111111 11

RESULT 11

US-08-222-177A-361/c
Sequence 361, Application US/08222177A
Patent No. 5582979

GENERAL INFORMATION:

APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:

INFORMATION FOR SEQ ID NO: 361:

SEQUENCE CHARACTERISTICS:
LENGTH: 49 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: m6114rs
US-08-222-177A-361

Query Match 0.5%; Score 26.8; DB 1; Length 49;
Best Local Similarity 73.9%; Pred. No. 2.9e+02;
Matches 34; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4078 TGTATGTGTGGCTGTGGTGTGGGTGGAGTGTGTCTGCTGAGTAA 4123
111 111111 1111 1111 111 1 111111 1 11 111

Db 47 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATGATGATTA 2

RESULT 12

US-08-222-177A-304/c
Sequence 304, Application US/08222177A
Patent No. 5582979

GENERAL INFORMATION:

APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison

STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 304:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd94rs
US-08-222-177A-304

Query Match 0.5%; Score 26.8; DB 1; Length 53;
Best Local Similarity 81.6%; Pred. No. 3.1e+02;
Matches 31; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4078 TGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4115
Db 45 TGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8

RESULT 13
US-08-222-177A-412/c
Sequence 412, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.

REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 412:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-222-177A-412

Query Match 0.5%; Score 26.8; DB 1; Length 56;
Best Local Similarity 81.6%; Pred. No. 3.2e+02;
Matches 31; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4078 TGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4115
Db 38 TGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1

RESULT 14
US-08-222-177A-134/c
Sequence 134, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd28rs
US-08-222-177A-134

Query Match 0.5%; Score 26.8; DB 1; Length 65;
Best Local Similarity 81.6%; Pred. No. 3.5e+02;

Matches 31; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4078 TGTATGTGGCTGTGGTGTGGTGGAGTGTCTG 4115
 ||| ||||| ||| ||||| ||| ||||| |||
 DB 57 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 20

RESULT 15

US-08-222-177A-107/c
 ; Sequence 107, Application US/08222177A
 ; Patent No. 5582979

GENERAL INFORMATION:

APPLICANT: Weber, James L.
 TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
 TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
 NUMBER OF SEQUENCES: 460
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Dewitt Ross & Stevens, S.C.
 STREET: 8000 Excelsior Drive, Suite 401
 CITY: Madison

STATE: Wisconsin
 COUNTRY: USA
 ZIP: 53717-1914

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/222,177A

FILING DATE:
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/341,562
 FILING DATE: 21-APR-1989

ATTORNEY/AGENT INFORMATION:

NAME: Sara, Charles S.
 REGISTRATION NUMBER: 30,492
 REFERENCE/DOCKET NUMBER: 09865.601
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (608) 831-2100
 TELEFAX: (608) 831-2106

TELEX:

INFORMATION FOR SEQ ID NO: 107:

SEQUENCE CHARACTERISTICS:

LENGTH: 91 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

IMMEDIATE SOURCE:

CLONE: mfd19rs

US-08-222-177A-107

Query Match 0.5%; Score 26.8; DB 1; Length 91;

Best Local Similarity 68.5%; Pred. No. 4.4e+02;

Matches 37; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 4062 AGGCTGAGGCGCCCATGTATGTGTGGCTGTGGTGTGGTGGAGTGTCTG 4115
 || ||||| ||| ||||| ||| ||||| ||| ||||| |||
 DB 81 AGAGTGAGTG 28

Search completed: November 24, 2002, 00:40:09
 Job time : 122 secs

